PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/87, 15/82, C12P 21/02, C12N 7/01 — A1

(11) International Publication Number: WO 97/39134

(43) International Publication Date: 23 October 1997 (23.10.97)

(21) International Application Number: PCT/GB97/01065

(22) International Filing Date: 17 April 1997 (17.04.97)

(30) Priority Data:
9607899.3 17 April 1996 (17.04.96) GB

(71) Applicant *(for all designated States except US)*: SCOTTISH CROP RESEARCH INSTITUTE [GB/GB]; Invergowrie, Dundee DD2 5DA (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: WILSON, Thomas, Michael, Aubrey [GB/GB]; The Coach House, 4 Balruddery Meadows, Invergowrie, Dundee DD2 5LJ (GB). CHAPMAN, Sean, Nicholas [GB/GB]; 47 Fort Street, Dundee DD5 2AB (GB).

(74) Agent: MURGITROYD & COMPANY; 373 Scotland Street, Glasgow G5 8QA (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: VIRUS-LIKE PARTICLE

(57) Abstract

Chimaeric pseudovirus particles and a method for producing a foreign protein using the same are disclosed. The pseudovirus particles comprise a protein (e.g. a coat protein) having a viral portion and a non-viral portion, and a nucleic acid (optionally chimaeric) to stabilize the aggregation of the protein, and create a helical ribonucleocapsid with the structure and symmetry approaching the native virus.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

Virus-like particle.

This invention relates to a virus-like particle, especially to a pseudovirus particle, and to a method for the production of a chimaeric protein using such virus-like particles. The protein can be a capsid protein which can self assemble *in vivo* with the nucleic acid (which may be chimaeric) to form the virus-like particles.

Pseudovirus particles are virus-like particles comprising viral coat protein subunits and a portion of the wild-type viral nucleic acid. Pseudoviruses may also include foreign nucleic acid. The coat protein can be wild-type, modified or chimaeric. A pseudovirus may lack at least a portion of the wild-type viral nucleic acid (or may possess a non-functional analogue of the wild-type nucleic acid) and this commonly renders the pseudovirus incapable of some function which is characteristic of the wild-type virus, such as replication. Alternatively or additionally, other genes may be missing or disabled, and the pseudovirus may be, for example, replication competent but incapable of cell-cell movement. The missing or dysfunctional gene(s) can be provided on the genome of a host cell or on a plasmid etc present in the host cell, thereby restoring the function of the wild-type virus to the pseudovirus when in the host cell.

The physical properties of the pseudovirus particle such as shape, symmetry, nucleic acid:protein ratio are usually similar to or identical with the wild-type

virus from which the pseudovirus is derived, although particle length and width can be influenced by nucleic acid length and coat protein composition respectively.

According to the present invention there is provided a virus-like particle comprising nucleic acid and protein, the protein having a first (viral) portion and a second (non-viral) portion.

The term "virus-like particle" refers to self-assembling particles which have a similar physical appearance to virus particles and includes pseudoviruses. Virus-like particles may lack or possess dysfunctional copies of certain genes of the wild-type virus, and this may result in the virus-like-particle being incapable of some function which is characteristic of the wild-type virus, such as replication and/or cell-cell movement.

The nucleic acid can be DNA or RNA, according to the genome of the virus from which the virus-like particle is derived. The nucleic acid may comprise an origin-of-assembly sequence (OAS) by which we mean a nucleic acid sequence which permits initiation of assembly of the protein and nucleic acid into virus-like particles.

Further according to the invention there is provided a method of producing a protein having a first (viral) portion and a second (non-viral) portion, the method comprising expressing the protein in a cell, providing a nucleic acid sequence capable of assembly with the protein into a virus-like particle (VLP), and permitting *in vivo* assembly of the protein and nucleic acid into VLPs.

The virus-like particles can be purified from the cell

by standard techniques such as centrifugation etc, and the chimaeric protein can optionally be cleaved to release the second portion from the first portion, or separated entirely from the nucleic acid. If the chimaeric protein is left attached to the virus-like particle, the whole virus-like particle can also be used for presentation of peptide epitopes for vaccination of animals, the production of therapeutic or industrial proteins and polypeptides and/or the delivery of therapeutic nucleic acid molecules (optionally targeted delivery), such as ss or ds DNA or RNA, including antisense molecules.

The nucleic acid can advantageously be provided from a plasmid in the cell, possibly by transcription of such a plasmid. The protein may be encoded by the same or another plasmid in the cell. Alternatively, one or both of the nucleic acid and protein can be coded from the genome of the cell.

The cell is preferably a bacterium such as *E. coli* although other forms of bacteria and other cells may be useful, such as mammalian cells, plant cells, yeast cells and insect cells. The cell may be a natural host cell for the virus from which the virus-like particle is derived, but this is not necessary.

The use of a cell for the assembly of the virus-like particle *in vivo* enables facile cell handling techniques to be employed to facilitate purification of virus-like particles and purification of protein. In addition, where it is desired to produce a second portion protein which is toxic to some cells (eg plant cells) a different (eg bacterial) cell may be employed.

The nucleic acid is preferably chosen in accordance

with its ability to assemble with the viral protein. For example, the virus-like particle may be derived from tobacco mosaic virus (TMV). In such a case, the first portion of the protein is preferably derived from TMV coat protein (CP), and the nucleic acid has at least an OAS of eg 75 or more nucleotides derived from TMV RNA. The sequence of the remainder of the nucleic acid is not important, and it can be chosen to code for the chimaeric protein or may be of some other eg unrelated or therapeutic sequence. The inclusion of nucleic acid in the virus-like particle means that the particle is of helical symmetry and more stable than simple aggregations of protein (eg planar, stacked or helical arrays), which are normally created at low pH *in vitro* from purified TMV coat protein, and can dissociate outside a narrow pH range. Also, the length of the particle can be selected by specifying a particular length of nucleic acid. This results in a more uniform range of particle sizes, which has advantages in purification procedures such as centrifugation, and in defining and regulating the quality control for products for medical use.

A further advantage with the use of nucleic acid in the assembly of virus-like particles is that the resultant particle can have a regular multivalent and true helical structure which can be more immunogenic than an aggregation of protein or free subunits of protein. The greater stability of the particle can also provide longer access to the immune system in certain embodiments.

The second portion of the chimaeric protein is preferably disposed on the outer surface of the virus-like particle. Thus where the particle is derived from TMV, the second portion can be disposed on the amino or

carboxy terminus, or inserted in eg an internal loop disposed on the outer surface of the CP. This can result in improved assembly as compared with the assembly of particles having a second portion on another location of the CP, and can enhance immune recognition of the second portion on the particle surface, which is useful for embodiments where the CP is an immunogen such as a vaccine. In certain cases it may be possible to provide large second portion proteins.

It is advantageous to use a virus which is flexuous (ie which can bend easily) since chimaeric proteins with large second portions may be able to assemble more easily into virus particles which are flexuous than those which are rigid. PVX is preferred since it forms a flexuous particle.

A linker peptide can be incorporated between the first and second portions and may have the function of spacing the two portions from one another, reducing steric restrictions. Optionally the linker peptide may contain a cleavage site.

The term "cleavage site" refers to a short sequence of amino acids which is recognisable and subsequently cleavable by eg a proteolytic enzyme or by chemical means. Suitable proteolytic enzymes include trypsin, pepsin, elastase, factor Xa etc. Alternatively the cleavage site may be vulnerable to cleavage by other means, for example by addition of chemicals such as cyanogen bromide (CNBr) or acids.

The term "cleavage site" may also include sequences that self-leave such as the FMDV (Foot and Mouth Disease Virus) 2A protease.

The cleavage site may be an integral part of either the first or second portion. Hence either/or both of the portions may include an integral cleavage site.

The second portion protein may be a short oligopeptide (10-40 amino acids) or a relatively large polypeptide eg over 10kDa. Proteins of 25-30 kDa may also be suitable for production by the method of the invention.

The first (viral) portion of the chimaeric protein may be any protein, polypeptide or parts thereof, derived from a viral source including any genetically modified versions thereof (such as deletions, insertions, amino acid replacements and the like). In certain embodiments the first portion will be derived from a viral coat protein (or a genetically modified version thereof). Mention may be made of the coat protein of Potato Virus X as being suitable for this purpose. Preferably the first portion has the ability to assemble into virus-like particles by first-portion/first portion association. Thus, a chimaeric protein molecule can assemble with other chimaeric protein molecules or with wild-type coat protein into a chimaeric virion.

In a preferred embodiment of the invention the particle is derived from a tobamovirus such as tobacco mild green mosaic virus TMGMV), tobacco mosaic virus (TMV), or from a potexvirus such as PVX, and in such an embodiment, the second portion is preferably disposed at or adjacent the N-terminus of the coat protein. In PVX, the N-terminus of the coat protein is believed to form a domain on the outside of the virion.

The second portion of the chimaeric protein may be any protein, polypeptide or parts thereof, including any

genetically modified versions thereof (such as deletions, insertions, amino acid replacements and the like) derived from a source other than the virus from which the first portion is derived. In certain embodiments the second portion or the protein derived therefrom is a biologically active or otherwise useful molecule. The second portion or the protein derived therefrom may also be a diagnostic reagent, an antibiotic or a therapeutic or pharmaceutically active agent. Alternatively the second portion protein may be a food supplement.

It is not necessary for the first portion to comprise a whole virus coat protein, but this remains an option. Some no-essential amino acids could be removed during construction of the CP gene.

The virus particle may be formed by the assembly of chimaeric proteins only or by the mixed assembly of chimaeric proteins together with some unmodified or less modified forms of the naturally occurring wild-type coat protein which forms the basis of the first portion. For a mixed virus particle of the latter type, there must be present polynucleotide(s) encoding the chimaeric protein and the naturally occurring coat protein. The appropriate protein-coding sequence(s) may be arranged in tandem on the same molecule, or could be generated by differential RNA splicing Alternatively, the different proteins could be translated from the same nucleotide sequence and modified later, eg by *in vivo* processing such as self cleavage. An example of this is the provision of a chimeric CP gene encoding eg GFP-2A-CP fusion protein, which is expressed from a single gene (eg on a plasmid, from the genome of the cell, or from the RNA of the VLP) and which self cleaves a variable number of the

translated proteins into separate GFP and CP, a proportion of the translated proteins remaining uncleaved as GFP-2A-CP. Thus a heterologous mixture of CPs can be assembled into a VLP, with eg every 10th CP bearing a second portion, and the remaining CPs being cleaved, native (or substantially native) CPs. Thus the potential problems with stearic hindrance which might occur if all the CPs were chimaeric can be overcome. Suitable co-translational cleavage sequences can be chosen for particular cell types. The efficiency of the co-translational cleavage can be modified to produce the required proportion of cleaved/whole CPs in the assembled VLP.

An advantage is gained by using a virus which forms a helical particle with a relatively large pitch. PVX has a pitch of 3.4nm and is to be preferred over viruses with a lower pitch. Virus particles with higher pitches may be able to accommodate larger protein insertions on their surfaces since their coat proteins assemble with more space between them than coat proteins of viruses with lower pitches.

The method can be used for expression of metabolic enzymes for pathway engineering, nutritional supplements (eg hi-met proteins), anti-potato cyst nematode lectins, gut protease inhibitors, anti-botrytis agents, PGIPs, anti-insect *Bacillus thuringiensis* toxin and herbicide resistance agents, industrial enzymes, pharmaceuticals, therapeutic proteins and nucleic acids, and as bioreactors.

While modifications and improvements may be incorporated without departing from the scope of the invention, embodiments will now be described by way of the following examples and with reference to the

accompanying drawings in which:

Fig 1 is a schematic representation of the plasmid pA27;
Fig 2 is an SDS PAGE analysis of proteins from purified TMV and pseudovirions. Samples were electrophoresed on an SDS/PAGE gel and silver stained. Lane 1, purified TMV. Lane 2, VLPs purified from E. coli BL21(DE3) cells transformed with plasmids pA27 and pLys102. The positions of coelectrophoresed marker proteins and their molecular weights in kDa are shown to the left;
Fig 3 is an electron microscope image of VLPs. VLPs purified from E. coli BL21(DE3) cells transformed with plasmids pA27 and pLys102 were negatively stained with 2% sodium phosphotungstate pH 5.0 and viewed in the electron microscope. Magnification x 20,000;
Fig 4 shows sequence information for LITMUS 39 plasmids used in Example 2;
Fig 5 shows a schematic representation of cDNA constructs used in Example 2;
Fig 6 shows immunoblot analysis of extracts of leaves probed with anti-CP antiserum; and
Fig 7 shows immunoblot analysis of virus prepared from plants infected with a VLP.

Example 1:

A sequence encoding two glycine residues and an eight amino acid antigenic epitope (EQPTTRAQ) from VP1 of poliovirus type 3 [1] was fused to the 3' end of a synthetic gene coding for the tobacco mosaic virus (TMV) coat protein by PCR amplification with mutagenic primers. The plasmid pTMVCP [1] was used as a template for amplification with primers P1311 (5' AAG-AAT-TCA-TAT-GTC-TTA-TTC-GAT-TAC-C 3') and P1312 (5' AAG-GAT-CCT-CAC-TGA-GCA-CGA-GTA-GTC-GGC-TGT-TCA-CCA-CCA-GTT-GCC-GGG-CCC-GAG 3'). The amplification product was

treated with T4 DNA polymerase to make it blunt-ended and ligated into *EcoRV* digested pKR [2]. The ligation products were transformed into *E. coli* strain JM101. Transformants were screened for the desired plasmid, pA11, containing the gene encoding the modified TMV coat protein.

To enable expression of the modified TMV coat protein in *E. coli* a fragment encompassing the modified gene was cloned into an expression vector, under the transcriptional control of T7 promoter and Tø terminator sequences. The plasmid pA11 was digested with *Nde*I and *Bam*HI and the 510 base pair fragment released was cloned between the same sites of pET3a [3]. The nucleotide sequence of the resulting plasmid, pA27 (Figure 1), in the region encoding the eight amino acid epitope and the linker of two glycine residues, was confirmed by nucleotide sequence determination.

In Figure 1, sequence encoding TMV coat protein and ten amino acid peptide fused to the carboxy-terminus are indicated by boxes marked TMV CP and PEP respectively. Restriction endonuclease sites used for the introduction of the modified TMV coat protein gene into the plasmid pET3a are indicated above. The T7 promoter and Tø terminator sequences from the plasmid pET3a are indicated by a double thickness arrow and line respectively. The nucleotide sequence of the 3' end of the modified TMV coat protein gene and the amino acids encoded by this sequence are shown below. The nucleotide sequence encoding the additional ten amino acids and the amino acids themselves are shown in bold.

To obtain expression of the modified TMV coat protein and production of pseudovirions the plasmid pA27 was transformed into *E. coli* BL21(DE3) cells that had

previously been transformed with the plasmid pLys102 [4]. The plasmid pLys102 produces a chimaeric RNA transcript encoding chloramphenicol acetyl transferase and containing the TMV origin-of-assembly sequence, which when co-synthesized with TMV coat protein in *E. coli* directs the assembly of pseudovirus particles of 70nm length (modal) and 18nm diameter. That plasmid pA27 directed the synthesis of modified TMV coat protein was confirmed by SDS/PAGE analysis of IPTG induced bacterial lysates [4]. Production of a TMV coat protein-related protein with a slightly lower mobility than unmodified TMV coat protein was detected by Coomassie blue staining and immunoblotting of SDS/PAGE gels as described by Hwang *et al.* [4].

VLPs containing the modified TMV coat protein were purified using a protocol based on that described by Hwang *et al.* [4]. Colonies of BL21(DE3) co-transformed with pA27 and pLys102 were used to inoculate 5 ml of M9ZB medium supplemented with 100 μg/ml ampicillin and 35 μg/ml chloramphenicol. Cultures were grown overnight at 37°C. The bacteria were pelleted from the overnight cultures and used to inoculate 500 ml of M9ZB medium supplemented with ampicillin and chloramphenicol. The large-scale cultures were grown at 37°C until mid-log phase (A_{600} = 0.7). Cultures were induced with 0.4 mM IPTG and incubated at 30°C for eighteen hours. Cells were harvested by centrifugation (4800 x *g*, 4°C, 6 min). Bacterial pellets were resuspended in 3ml of TE (10 mM Tris-HCl pH 7.5 / 1 mM EDTA) and incubated with lysozyme (0.4 mg/ml) at 20°C for 60 min. Bacteria were lysed by addition of 4 ml 40% w/v sucrose in TE and then 16 ml of TE. DNase I was added to 6.5 μg/ml and the lysates incubated at 37°C for 90 min. Bacterial debris was removed by centrifugation (20800 x *g*, 4°C, 30 min). The resulting supernatants were extracted with

10 ml of chloroform and the two phases separated by centrifugation (9200 x *g*, 4°C, 10 min). 3.7 ml of 5M NaCl and 2.63 ml of 40% polyethylene glycol (average molecular weight 6000) were added to 20 ml of the aqueous phase. The solutions were mixed and incubated on ice for 60 min. Precipitated material was collected by centrifugation (20800 x *g*, 4°C, 15 min). The pelleted material was resuspended in 1 ml of TE. Insoluble material was removed by centrifugation (16000 x *g*, 4°C, 5 min). The supernatant was centrifuged (160000 x *g*, 4°C, 120 min) on a sucrose gradient (10-40% w/v in TE). Fractions were collected from the gradients and those containing helical TMV-like particles, assessed by double-antibody sandwich ELISA with a mouse monoclonal antibody specific for an epitope in the TMV coat protein helix as described by Hwang *et al.* [4], were pooled for further purification.

VLPs were collected by centrifugation (235,000 x *g*, 15°C, 150 min). Pelleted pseudovirions were resuspended in 0.5 ml of TE. Insoluble material was removed by centrifugation (840 x *g*, 4°C, 5 min). The supernatant was centrifuged (189,000 x *g*, 15°C, 120 min) on a CsCl gradient (10-40% (wt/wt) in TE). Bands containing pseudovirus were collected from the gradients and dialyzed against 50 mM sodium phosphate pH 7.0.

The yield of VLPs was estimated by measuring the absorption at 260 nm. The final yield of pseudovirus was 5.8 mg from 500 ml of culture. The purity of the pseudovirus preps was assessed by silver staining of samples electrophoresed on SDS/PAGE gels (Figure 2). On SDS/PAGE gels the unmodified TMV coat protein produced by pET302 and the modified coat protein produced by pA27 migrate relative to protein standards (Bio-Rad) with apparent molecular weights of 20.9 kDa and 22.6

kDa respectively. The predicted molecular weights for these two proteins are 17.7 kDa and 18.6 kDa respectively.

The integrity of the pseudovirus preparations was assessed by negative staining of pseudovirus samples with 2% sodium phosphotungstate and observation of the stained samples in the electron microscope (Figure 3). Pseudovirus preparations were diluted to 1 mg / ml in 50 mM sodium phosphate pH 7.0 for immunization of mice.

Example 2:

A plasmid containing the tobacco mild green mosaic virus (TMGMV) coat protein (CP) gene and 3' untranslated region (UTR) was produced to facilitate the production of green fluorescent protein (GFP), foot-and-mouth disease virus 2A, TMGMV CP gene fusions. A 955 base pair (bp) fragment containing the TMGMV CP and 3' UTR was PCR amplified from the plasmid 30B (W.O. Dawson, Citrus Research and Education Center) using the primers TMGMV-CP-Apa (5' CAA-TGG-GCC-CTA-TAC-AAT-CAA-CTC-T 3') and M13-Reverse (5' AGC-GGA-TAA-CAA-TTT-CAC-ACA-GGA 3'). The primer TMGMV-CP-Apa was designed to mutagenize the sequence coding for the initiating methionine and first proline codon of the TMGMV CP to an ApaI restriction enzyme site. This results in the conversion of the methionine codon to a glycine codon, but maintains the proline codon. The 837bp fragment released by digestion of the PCR amplification product with the restriction endonucleases ApaI and KpnI was cloned into the 3322bp fragment released by digestion of pSL1180 (Pharmacia) digested with the same restriction endonucleases and treated with calf intestinal alkaline phosphatase. The resulting plasmid was named pSL.TMGMV-CP-UTR.

CFP-2A-TMGMV CP gene fusions were produced by cloning DNA fragments containing GFP-2A fusions into pSL.TMGMV-CP-UTR adjacent to the codon for the first proline in the TMGMV CP gene. A selection of LITMUS 39 (New England Biolabs) based plasmids containing GFP-2A-potato virus X CP gene fusions were used as sources for the GFP-2A gene fusion.

The nucleotide sequence and amino acids encoded by the different LITMUS 39 based plasmids between the carboxy-terminal lysine codon of the GFP gene and the amino-terminal proline codon of the PVX CP gene are shown in Figure 4.

These plasmids contain a variety of sequences coding for different 2A amino acid sequences between the carboxy-terminal lysine codon of GFP and the first proline codon of the PVX CP. Fragments of between 900 and 1050bp were PCR amplified from the plasmids pLit, GFP-$2A_{16H}$-CP, pLit.GFP-$2A_{16K}$-CP, pLit.GFP-$2A_{23H}$-CP and pLit.GFP-$2A_{58K}$-CP using the primers GFP-5'-Sal (5' TCA-ATC-GTC-GAC-ATG-AGT-AAA-GGA-GAA-GAA 3') and N3#4 (5' TGT-ACT-AAA-GAA-ATC-CCC-ATC-C 3'). The primer GFP-5'-Sal introduces a SalI restriction enzyme site upstream of the initiating methionine codon of the GFP gene. Fragments containing the GFP gene fused to the different 2A sequences were released by digestion of the PCR amplification products with SalI and ApaI and ligated into the large fragment released by digestion of pSL.TMGMV-CP-UTR with the same restriction enzymes and treated with calf intestinal phosphatase. The resulting plasmids were digested with SalI and BstEII and the released fragments containing the GFP-2A-TMGMV CP gene fusion and TMGMV UTR were introduced into the plasmid 30B digested with XhoI and BstEII to regenerate full-length TMV based clones. Thus the final clones

comprise wild-type TMV strain U1 sequence up to position 5757 in the CP gene, with the exception of a mutagenized CP initiating methionine codon, followed by a short polylinker sequence, the GFP-2A-TMGMV CP gene fusions and the TMGMV 3' UTR.

Figure 5 shows a schematic representation of viral cDNA constructs used in this example. Boxes represent coding sequences. The genes for the three viral proteins common to all constructs are indicated by their predicted Mr values (K=kDa). The genes for the green fluorescent protein, 2A oligopeptide and TMGMV CP are indicated by GFP, 2A and CP respectively. Restriction enzyme sites used in the cloning procedures are indicated above.

In vitro run-off transcripts were synthesized from KpnI linearized plasmids p30B.GFP-$2A_{16H}$-CP, p30B.GFP-$2A_{16K}$-CP, p30B.GFP-$2A_{23H}$-CP, p30B.GFP-$2A_{58K}$-CP and p30B.GFP, a derivative of p30B that has had the GFP gene introduced into the unique XhoI site of P30B, which expresses free GFP. The transcripts derived from all the plasmids were infectious when inoculated onto *Nicotiana benthamiana* plants; virus derived from transcript-infected plants is referred to subsequently by the name of the progenitor plasmid without the "p" prefix. Following inoculation, all the viruses caused the development of fluorescent regions which were first detectable by eye under UV illumination between three and four days post inoculation. Subsequent long distance movement of the virus led to the appearance of green fluorescence in systemically infected leaves. The appearance of fluorescence in systemically infected leaves occurred at a similar time, nine days post inoculation, for plants infected with 30B.GFP, 30B.GFP-$2A_{16H}$-CP and 30B.GFP-$2A_{16K}$-CP, but was delayed for

30B.GFP-2A$_{23H}$-CP and 30B.GFP-2A$_{58K}$-CP.

Western blotting of protein extracts from systemically infected *N. benthamiana* leaves, probed with rabbit polyclonal antisera raised against TMV CP (Figure 6), detected two protein species in each of the 30B.GFP-2A-CP infected samples. This result indicated that the modified viruses were producing a GFP-2A-CP fusion protein, the *in vivo* processing of which resulted in the production of a GFP-2A fusion protein and free TMGMV CP. For 30B.GFP-2A$_{16H}$-CP, 30B.GFP-2A$_{16K}$-CP and 30B.GFP-2A$_{58K}$-CP the majority of CP related protein produced was in the unfused form. Protein was prepared from mock-inoculated control plants (lane 1) or from plants inoculated with in vitro transcripts synthesized from plasmid DNAs (p30B.GFP, lane 2; p30B.GFP-2A$_{23H}$-CP, lane 3; p30B.GFP2A$_{16H}$-CP, lane 4; p30B.GFP-2A$_{16K}$-CP, lane 5; p30B.GFP-2A$_{58K}$-CP, Lane 6). Lane 7 contains 125ng of TMGMV CP. The predicted Mr values of TMGMV CP, GFP and GFP-2A-CPs are 17.5 kDa, 26.9 kDa and between 46 and 52 kDa, respectively. The Mr values of standards (X10$_{-3}$) are shown on the left.

The observation that the modified viral constructs were capable of rapid systemic movement like 30B.GFP suggested that they were also capable of virus particle formation. To confirm that this was the case homogenates were prepared by grinding fluorescent inoculated leaf tissue from plants infected with 30B.GFP and 30B.GFP-2A$_{23H}$-CP in a "mini-mortar" with 50mM phosphate buffer pH 6.5. The homogenates were applied to a carbon coated grid and stained with 2% sodium phosphotungstate pH 6.5 prior to observation in the electron microscope. 30B.GFP-2A$_{23H}$-CP was found to produce rod-shaped particles like those produced by 30B.GFP. To test whether the particles produced by

30B.GFP-$2A_{23B}$-CP had incorporated GFP-2A-CP fusion protein as well as free TMGMV CP immunotrapping (Roberts 1986, in Electron microscopy of proteins, Academic Press) was performed with rabbit polyclonal antisera raised against GFP and TMV.CP. While 30B.GFP infected tissue showed enhanced trapping with the TMV-CP antisera, but not with the GFP antisera, 30B.GFP-$2A_{23B}$-CP infected tissue showed enhanced trapping with both antisera (Table 1). This result suggested that the modified virus was capable of incorporating GFP-2A-CP fusion protein into particles.

Table 1

	Number of particles/1000μm²	
Coating antiserum	30B.GFP	30B.GFP-$2A_{23B}$-CP
None	223 +/- 57.0	3.5 +/- 1.33
TMV CP	4690 +/- 1200	58.0 +/- 3.16
GFP	112 +/- 9.45	67.5 +/- 15.2

To confirm this a virion extraction (Kearney *et al*, in Plant Molecular Biology Manual L1:1-16, Kluwer Academic Publishers) was performed on fluorescent, systemically infected tissue of plants infected with 30B.GFP-$2A_{16B}$-CP. Western blot analysis (Fig 7) of the virus preparation with GFP (B) and TMV CP (A) antisera demonstrated that the virus contained TMGMV CP and CGP-2A-CP fusion protein but no GFP-2A fusion protein. Mr values shown on left of Fig 7 ($x10^{-3}$) Thus the GFP-2A-CP fusion protein was assembled with free TMGMV CP into virus particles.

Modifications and improvements can be incorporated without departing from the scope of the invention.

Documents incorporated herein by reference:

[1]. Haynes, J.R., Cunningham, J., von Seefried, A., Lennick, M., Garvin, R.T. and Shen, S.-H. (1986). Bio/Technology, 4, 637-641. EP 0174759 A1 (Connaught Laboratories Limited), see particularly construction of plasmids.

[2] Waye, M.M.Y., Verhoeyen, M.E., Jones, P.T. and Winter, G. (1985). Nucleic Acids Research, 13, 8561-8571.

[3] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990). Methods in Enzymology, 185, 60-89.

[4] Hwang, D.-J., Roberts, I.M. and Wilson, T.M.A. (1994). Proceedings National Academy of Sciences U.S.A., 91, 9067-9071. WO 94/10329 (Rutgers University), see particularly deposit information therein.

Claims:

1 A method of producing a protein having a first (viral) portion and a second (non-viral) portion, the method comprising expressing the protein in a cell, providing a nucleic acid sequence capable of assembly with the protein into a virus-like particle (VLP), and permitting *in vivo* assembly of the protein and nucleic acid into VLPs.

2 A method as claimed in claim 1, wherein the VLPs are subsequently purified from the cell.

3 A method as claimed in claim 1 or claim 2, wherein after assembly the protein is cleaved to release the second portion from the first portion, or is separated entirely from the nucleic acid.

4 A method as claimed in any preceding claim, wherein the nucleic acid is provided from a plasmid.

5 A method as claimed in claim 4, wherein the protein is encoded by the same or another plasmid in the cell, or from the genome of the cell.

6 A method as claimed in any preceding claim, wherein the cell is selected from bacterial cells, mammalian cells, plant cells, yeast cells and insect cells.

7 A method as claimed in claim 6, wherein the cell is a natural host cell for the virus from which the virus-like particle is derived.

8 A method as claimed in any preceding claim, wherein the second portion of the protein is disposed

on the outer surface of the VLP.

9 A method as claimed in any preceding claim, wherein the VLP is flexuous.

10 A method as claimed in any preceding Claim wherein a cleavage site is incorporated on one of, or between, said first and second portions.

11 A method as claimed in any preceding claim, wherein a linker peptide is incorporated between the first and second portions.

12 A method as claimed in any preceding claim, wherein the second portion has a molecular weight of up to 10 kDa.

13 A method as claimed in any one of claims 1 to 11, wherein the second portion has a molecular weight of between 10 kDa and 30 kDa.

14 A method as claimed in any one of claims 1 to 11, wherein the second portion has a molecular weight over 30kDa.

15 A method as claimed in any preceding claim, wherein the first portion is derived from a viral coat protein or a modified version thereof.

16 A method as claimed in any preceding claim, wherein the first portion is derived from a tobamovirus or a potexvirus.

17 A method as claimed in any preceding claim, wherein the second portion or the protein derived therefrom is a biologically or pharmaceutically active

or useful molecule.

18 A method as claimed in any one of claims 1 to 16, wherein the second portion or the protein derived therefrom is a diagnostic reagent.

19 A method as claimed in any one of claims 1 to 16, wherein the second portion or the protein derived therefrom is a food supplement.

20 A method as claimed in any preceding claim, wherein the virus particle is formed by a mixed assembly of chimaeric proteins together with some unmodified or less modified forms of the naturally occurring wild-type coat protein which forms the basis of the first portion.

21 A method as claimed in claim 20, wherein the chimaeric proteins and the unmodified or less modified forms of the naturally occurring wild-type protein are expressed from different sequences of nucleic acid.

22 A method as claimed in claim 21, wherein the different sequences are on the same piece of nucleic acid in the cell.

23 A method as claimed in claim 21, wherein the different sequences are on different pieces of nucleic acid in the cell.

24 A method as claimed in claim 20, wherein the chimaeric proteins and the unmodified or less modified forms of the naturally occurring wild-type protein are expressed from the same sequence of nucleic acid.

25 A method as claimed in claim 24, wherein the

chimaeric proteins and the unmodified or less modified forms of the naturally occurring wild-type protein are generated by co-translational modification, or are modified after translation.

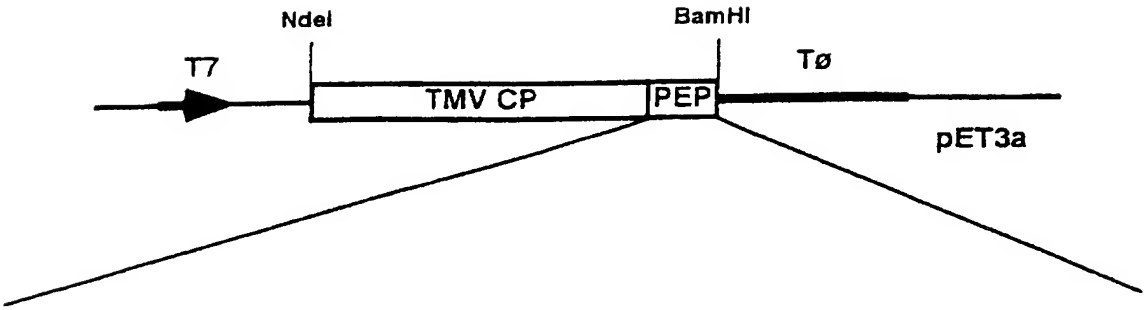
26 A method as claimed in any preceding claim, wherein the virus from which the first portion is derived forms a particle with a relatively high pitch of helix.

27 A virus-like particle (VLP) comprising nucleic acid and protein, the protein having a first (viral) portion and a second (non-viral) portion.

28 A VLP as claimed in claim 27 wherein the nucleic acid comprises an origin of assembly sequence which permits initiation of assembly of the protein and nucleic acid into VLPs.

29 A VLP as claimed in either of claims 27 or 28, wherein the second portion (non-viral) of the protein is disposed on the outer surface of the VLP.

FIG. 1

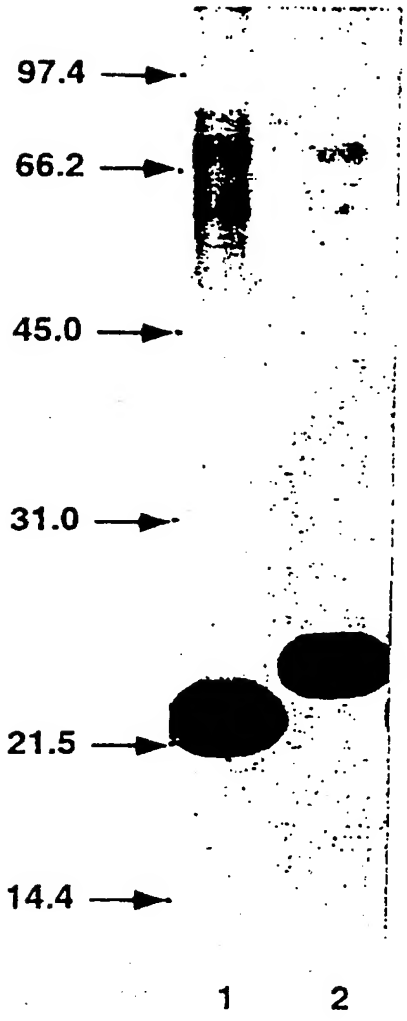


CCG-GCA-ACT-GGT-GGT-GAA-CAG-CCG-ACT-ACT-CGT-GCT-CAG-TGA-GGA-TCC

pro ala thr gly gly glu gln pro thr thr arg ala gln OPA

SUBSTITUTE SHEET (RULE 26)

FIG. 2



SUBSTITUTE SHEET (RULE 26)

FIG. 3

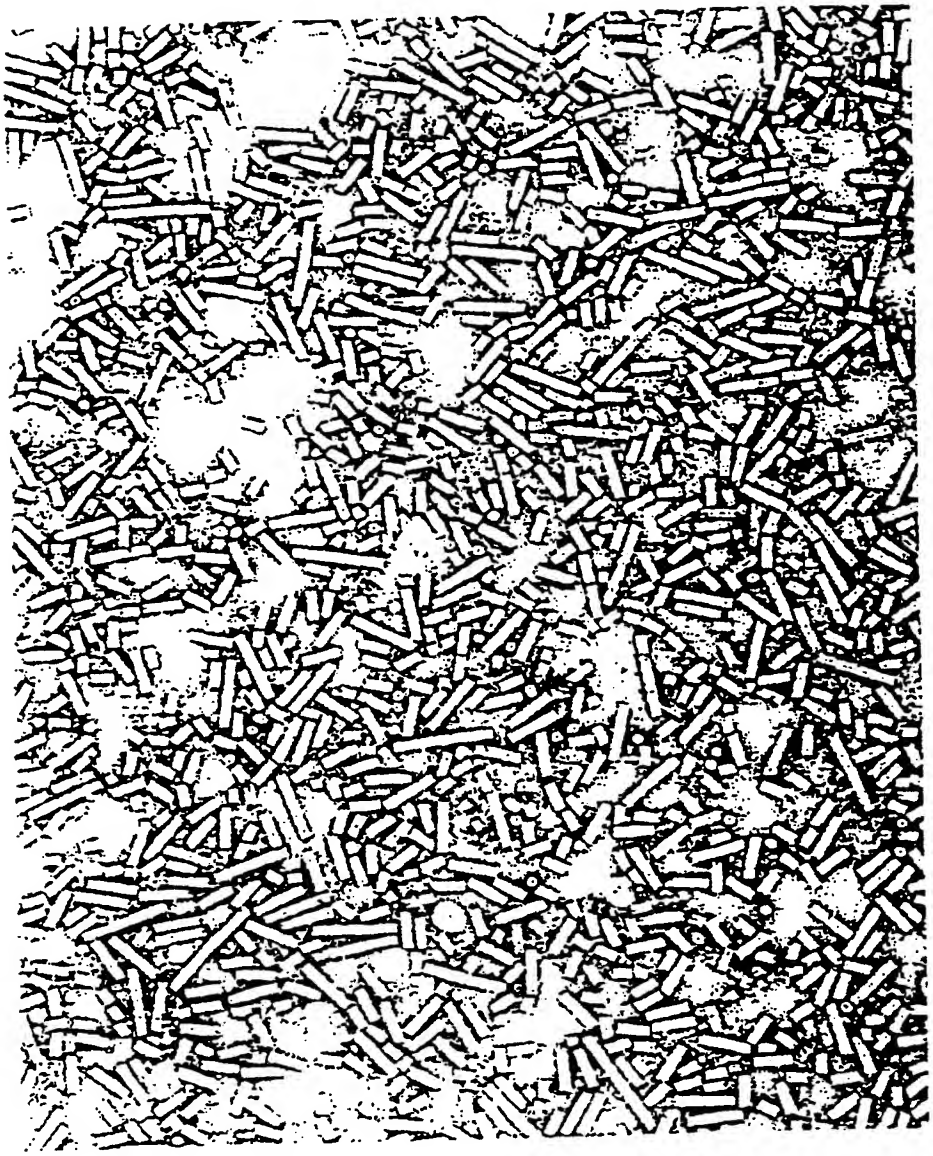


Figure 4.

pLit.GFP-2A$_{23H}$-CP

TCC GGA TCT AGA GCA CCT GTG AAA CAG CTG TTG AAT TTT GAC CTT CTT AAG CTT GCG GGA GAC GTC GAG TCC AAC CCT GGG
S G S R A P V K Q L L N F D L L K L A G D V E S N P G.

pLit.GFP-2A$_{16H}$-CP

TCC GGA TCT AGA AAT TTT GAC CTT CTT AAG CTT GCG GGA GAC GTC GAG TCC AAC CCT GGG
S G S R N F D L L K L A G D V E S N P G.

pLit.GFP-2A$_{16R}$-CP

TCC GGA TCT AGA AAT TTT GAC CTT CTC AAG TTG GCG GGA GAC GTC GAG TCC AAC CCT GGG
S G S R N F D L L K L A G D V E S N P G.

pLit.GFP-2A$_{SOK}$-CP

TCC GGA TCT AGA GTC ACC GAG TTG CTT TAC CGG ATG AAG AGG GCC GAA ACA TAC TGT CCA AGG CCC TTG CTG GCA ATC CAC CCA ACT GAA GCC
S G S R V T E L L Y R M K R A E T Y C P R P L L A I H P T E A.

AGA CAC AAA CAG AAA ATT GTG GCA CCG GTG AAA CAG ACT TTG AAT TTT GAC CTT CTC AAG TTG GCG GGA GAC GTC GAG TCC AAC CCT GGG
R H K Q K I V A P V K Q T L N F D L L K L A G D V E S N P G.

Figure 5.

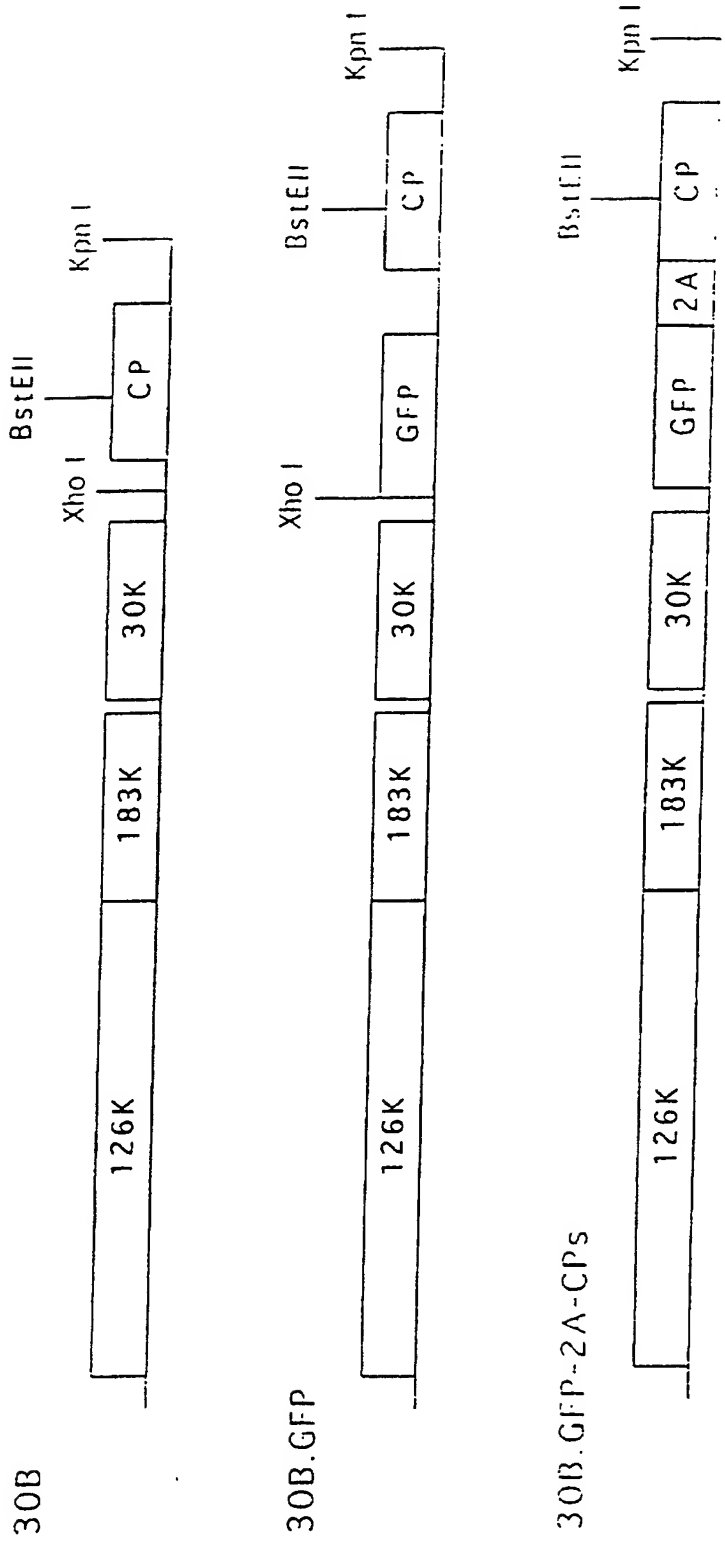


Figure 6.

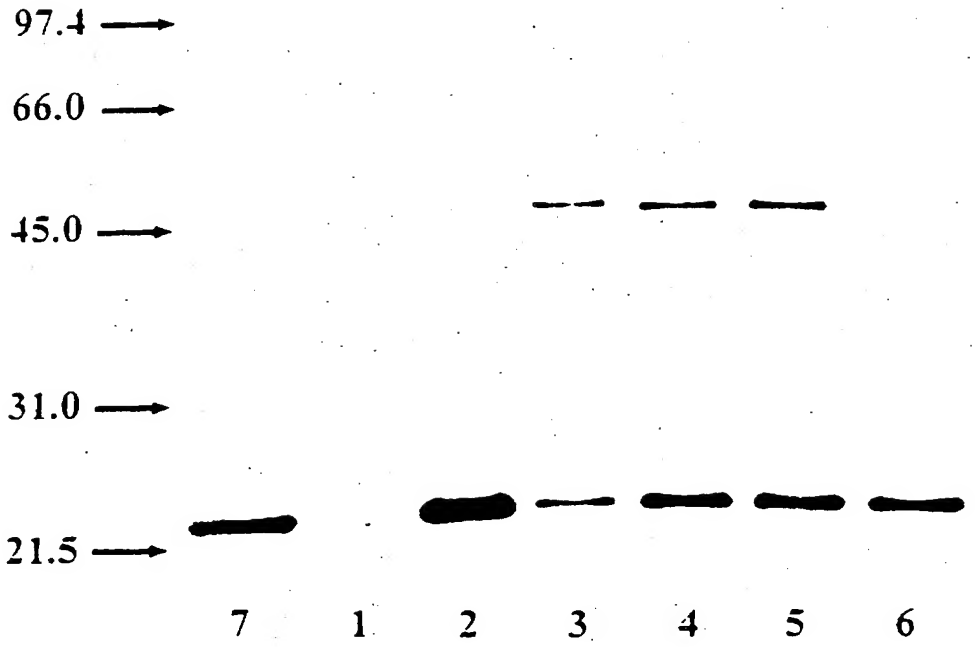
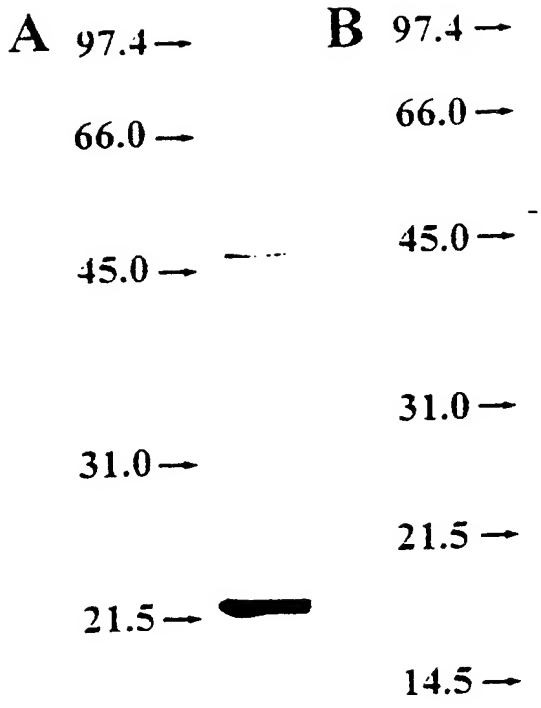


Figure 7.



SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

Internat Application No
PCT/GB 97/01065

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/87 C12N15/82 C12P21/02 C12N7/01

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	M.N. JAGADISH ET AL.: "High level production of hybrid Potyvirus-like particles carrying repetitive copies of foreign antigens in Escherichia coli" BIO/TECHNOLOGY., vol. 11, no. 10, October 1993, NEW YORK US, pages 1166-1170, XP002040652 see the whole document ---	1-6, 8-12, 15-18, 20-29
X	WO 96 05292 A (CONNAUGHT LABORATORIES LIMITED) 22 February 1996 see page 7, line 25 - page 12, line 27; figures 1-16 ---	1-12,15, 17,20-29
X	WO 95 10624 A (BOEHRINGER INGELHEIM INTERNATIONAL GMBH) 20 April 1995 see the whole document -----	1-12,15, 17,20-29

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
15 September 1997	23.09.97
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Authorized officer Cupido, M

Form PCT/ISA/210 (second sheet) (July 1992)

INTERN. IONAL SEARCH REPORT

Information on patent family members

In. ..ion Application No
PCT/GB 97/01065

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9605292 A	22-02-96	AU 3159995 A	07-03-96
		CA 2197446 A	22-02-96
		EP 0778888 A	18-06-97
WO 9510624 A	20-04-95	DE 4335025 A	20-04-95
		AU 7812094 A	04-05-95
		EP 0724643 A	07-08-96
		JP 9503665 T	15-04-97